Researchers create first image-recognition software that greatly improves web searches

November 18 2014

Dartmouth researchers and their colleagues have created an artificial intelligence software that uses photos to locate documents on the Internet with far greater accuracy than ever before.

The new system, which was tested on photos and is now being applied to videos, shows for the first time that a machine learning algorithm for image recognition and retrieval is accurate and efficient enough to improve large-scale document searches online. The system uses pixel data in images and potentially video - rather than just text—to locate documents. It learns to recognize the pixels associated with a search phrase by studying the results from text-based image search engines. The knowledge gleaned from those results can then be applied to other photos without tags or captions, making for more accurate document search results.

The findings appear in the journal *PAMI* (*IEEE Transactions on Pattern Analysis and Machine Intelligence*).

"Images abound on the Internet and our approach means they'll no longer be ignored during document retrieval," says Associate Professor Lorenzo Torresani, a co-author of the study. "Over the last 30 years, the Web has evolved from a small collection of mostly text documents to a modern, gigantic, fast-growing multimedia dataset, where nearly every page includes multiple pictures or videos. When a person looks at a Web page, she immediately gets the gist of it by looking at the pictures in it. Yet, surprisingly, all existing popular search engines, such as Google or

Bing, strip away the information contained in the photos and use exclusively the text of Web pages to perform the document retrieval. Our study is the first to show that modern machine vision systems are accurate and efficient enough to make effective use of the information contained in image pixels to improve document search."

The researchers designed and tested a machine vision system - a type of artificial intelligence that allows computers to learn without being explicitly programmed—that extracts semantic information from the pixels of photos in Web pages. This information is used to enrich the description of the HTML page used by search engines for document retrieval. The researchers tested their approach using more than 600 search queries on a database of 50 million Web pages. They selected the text-retrieval search engine with the best performance and modified it to make use of the additional semantic information extracted by their method from the pictures of the Web pages. They found that this produced a 30 percent improvement in precision over the original search engine purely based on text. The new system was developed by researchers at Dartmouth College, Tecnalia Research & Innovation and Microsoft Research Cambridge.

Provided by Dartmouth College